KONINKL. NEDERL. AKADEMIE VAN WETENSCHAPPEN – AMSTERDAM
Reprinted from Proceedings, Series C, 55, No. 5, 1952

BIOLOGY

THE NUTRITIVE VALUE OF BUTTERFAT, COMPARED WITH THAT OF VEGETABLE FATS

II. STATISTICAL ANALYSIS

BY

C. VAN EEDEN

(*Report S 63 of the Statistical Department of the Mathematical Centre at Amsterdam*)

(Communicated by Prof. D. VAN DANTZIG at the meeting of September 27, 1952)

1. Introduction

J. Boer and B. C. P. Jansen and afterwards E. H. Groot and C. Nieman executed experiments about the nutritive value of butter in comparison with that of vegetable fats to which vitamins A and D were added [1]).

The animals used were rats, mostly of the male sex; these rats were brought into the experiments 4 weeks after birth and fed with different diets. In all cases the increase in weight of the rats after 7 weeks and sometimes after 12 weeks is given.

2. Experiments of Boer and Jansen

2. 1. *The diets*

In these experiments three different diets were applied:

A: Summerbutter

B: Vegetable fat, with addition of 70 γ carotene and 0.35 γ calciferol a week

C: Vegetable fat, with addition of 70 γ carotene and 0.013 γ calciferol a week.

2. 2. *Testing the homogeneity of the subgroups*

First of all the question arises whether the experiments with one single diet in different seasons [2]) may be considered as samples from a single normal population.

Three different tests are applied, which have very different power functions:

1st. The test of normality of Geary and Pearson (1938). This test will lead to rejection of the hypothesis tested if the group of all of the observations taken together deviates strongly in form from the normal distribution.

The test is based on the statistics:

$$a = n^{-\frac{1}{2}} \cdot \{\sum_{i=1}^{n} (x_i - \bar{x})^2\}^{-\frac{1}{2}} \cdot \sum_{i=1}^{n} |x_i - \bar{x}|$$

[1]) Cf. part I of this paper, these Proceedings 55, 587–604; the observations are given there.

[2]) A group of rats with one diet in one season is called a *subgroup*.

and

$$b_1^{\frac{1}{2}} = n^{\frac{1}{2}} \cdot \{\sum_{i=1}^{n} (x_i - \bar{x})^2\}^{-3/2} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^3,$$

where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

2nd. The test of BARTLETT–HARTLEY (see BARTLETT 1934, HARTLEY 1940 and HARTLEY and PEARSON 1946). This test is especially powerful with respect to alternatives with inequality of the variances of the subgroups. It is therefore a test for the hypothesis that the variances of the subgroups are equal. The test is based on the ratio of the geometrical and the arithmetical mean of the variances of the subgroups.

3rd. The F-test of FISHER–SNEDECOR used in the analysis of variance with one classification (see e.g. M. G. KENDALL 1947, Vol. II, Chapter 23). This test is especially powerful with respect to inequality of the means of the subgroups and is therefore mainly a test for the hypothesis, that the means of the subgroups are equal.

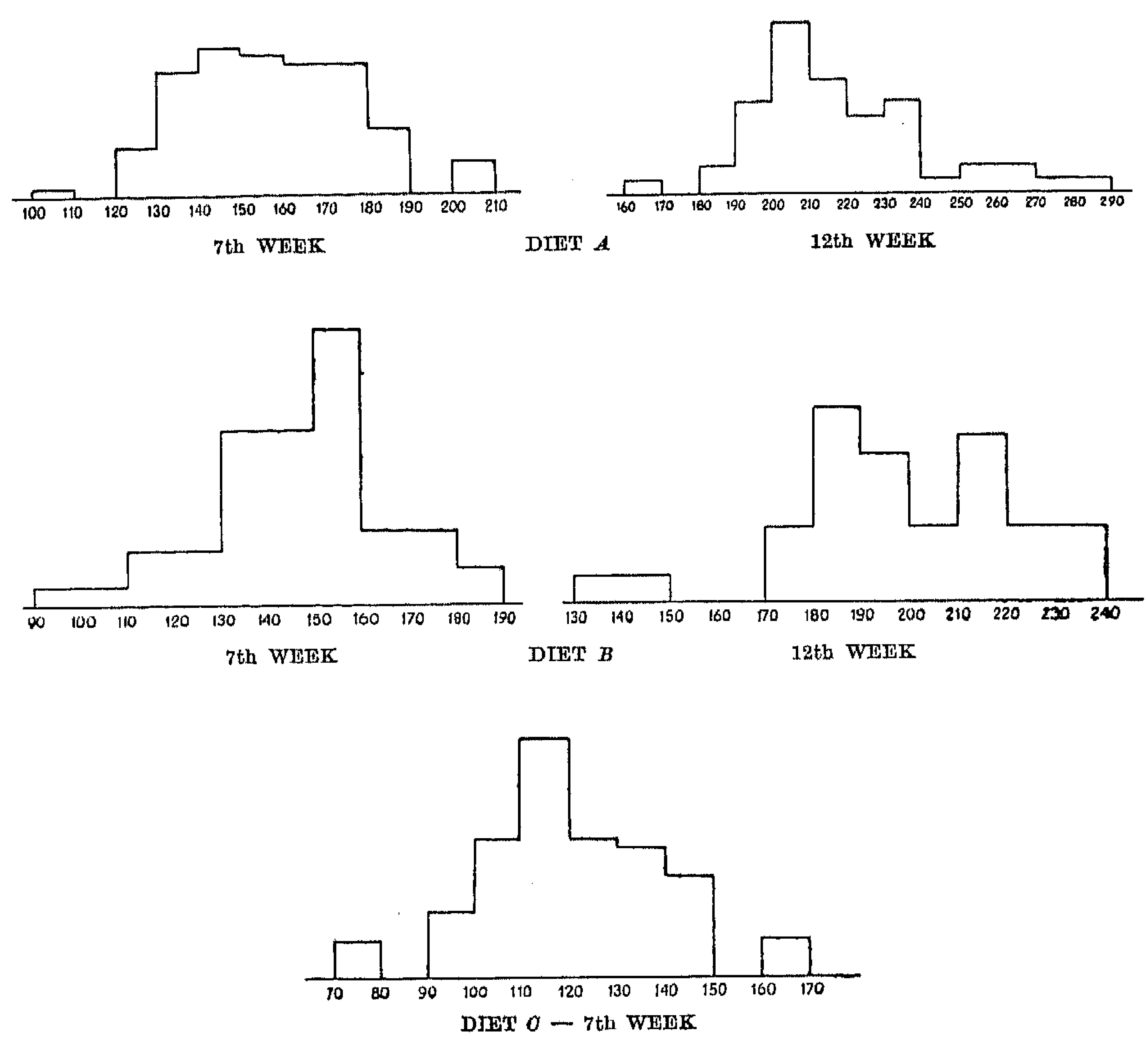

Fig. 1

The results are found in the tables 1, 2 and 3. The test of normality is only applied to the diets A (after 7 and 12 weeks) and B (after 7 weeks), which, according to figure 1, where we find histograms for the different diets, show the largest deviations from normality.

TABLE 1

Test of normality for some diets separately:

Diet	Growth after	Number of observations	a	Bilateral tail-probability [3]	$b_1^{\frac{1}{2}}$	Bilateral tail-probability [3]
A	7 weeks	83	0.80	1.00	0.3	0.40
A	12 ,,	45	0.79	0.80	0.6	0.08
B	7 ,,	48	0.78	0.55	−0.3	0.50

There proves to be no reason to reject the hypothesis of normality.

TABLE 2

Test of Bartlett-Hartley for each diet separately:

Diet	Growth after	Number of subgroups	M	Tail-probability [3]
A	7 weeks	11	8.9	> 0.05
A	12 ,,	7	6.4	> 0.05
B	7 ,,	5	1.1	$\gg 0.05$
B	12 ,,	4	7.5	> 0.05
C	7 ,,	4	1.8	$\gg 0.05$

Conclusion: there is no reason to reject the equality of the variances of the subgroups.

TABLE 3

Analysis of variance for each diet separately:

Diet	Growth after	z	Tail-probability [3]
A	7 weeks	0.18	0.40
A	12 ,,	0.09	> 0.40
B	7 ,,	−0.10	> 0.40
B	12 ,,	−0.23	> 0.40
C	7 ,,	0.11	> 0.40

Again there is no reason to reject the hypothesis tested, which is in this case the equality of the means of the subgroups.

From these non-rejections we cannot conclude that there are really no deviations from normality at all, but these results provide a support for the opinion, that such deviations, if present, are no serious objection against combining the subgroups of every diet to form dietgroups.

[3] The tail-probability of a test, with respect to a given observational result, is the level of significance of the smallest critical region of this test method, which contains this result.

2. 3. *Testing for differences between the dietgroups*

To investigate the difference in nutritive value between the diets STUDENT's test (see e.g. M. G. KENDALL 1947, Vol. II, p. 109) will now be applied to pairs of dietgroups, after testing the assumptions of normality and equality of the variances of the groups to be compared, which are inherent to STUDENT's test.

The assumption of normality has already been tested for every dietgroup separately (see § 2. 2: 1st). The assumption of equality of the variances is now tested for the pairs of dietgroups by means of the F-test of FISHER–SNEDECOR (cf. e.g. P. G. HOEL, 1948, p. 153). This test is based on the ratio of the variances of the samples. In table 4 we find the results:

TABLE 4

F-test for all pairs of dietgroups:

Diet	Growth after	s^2	ν [4])	F	Bilateral tail-probability
A *B*	7 weeks 7 weeks	390 412	82 47	1.06	> 0.40
A *B*	12 weeks 12 weeks	622 543	44 34	1.15	> 0.40
A *C*	7 weeks 7 weeks	390 342	82 26	1.14	> 0.40
B *C*	7 weeks 7 weeks	412 342	47 26	1.20	> 0.40

There is no reason to reject the hypothesis of equality of the variances and we now proceed to apply STUDENT's test, which is based on the statistic:

$$t = (\bar{x} - \bar{y}) \{s_1^2 + s_2^2\}^{-\frac{1}{2}} \left\{\frac{n_1 n_2 (n_1+n_2-2)}{n_1+n_2}\right\}^{\frac{1}{2}},$$

where $\bar{x}$ and $\bar{y}$ represent the sample-means, n_1 and n_2 the numbers of observations in the two samples, and where

$$s_1^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \quad \text{and} \quad s_2^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2.$$

The results are given in table 5. (See following page).

The experiments of BOER and JANSEN thus lead to the conclusion, that there is a difference between the nutritive value of

1. summerbutter and vegetable fat to which 70 γ carotene and 0.35 γ calciferol a week is added;
2. summerbutter and vegetable fat to which 70 γ carotene and 0.013 γ calciferol a week is added;
3. vegetable fat with addition of 70 γ carotene and 0.35 γ calciferol a week and vegetable fat with addition of 70 γ carotene and 0.013 γ calciferol a week.

[4]) Numbers of degrees of freedom.

In all these cases the first mentioned diet has given the larger increase in weight.

TABLE 5
Student's test for pairs of dietgroups:

Diet	Growth after	$\bar{x}$ resp. $\bar{y}$	t	ν [5])	Bilateral tail-probability
A *B*	7 weeks 7 weeks	157 146	2.97	129	0.003
A *B*	12 weeks 12 weeks	218 198	3.58	78	$< 10^{-3}$
A *C*	7 weeks 7 weeks	157 120	8.50	108	$< 10^{-3}$
B *C*	7 weeks 7 weeks	146 120	5.59	73	$< 10^{-3}$

2. 4. *Comparison of subgroup XII with subgroups I—XI*

Apart from the data of the subgroups I—XI of the summerbutter diet, there were data available of a twelfth subgroup with this diet, which however, originated from an experiment of a later date. The question arises, whether this subgroup should be added to this dietgroup or not. To answer this question the F-test and STUDENT's test, mentioned in the foregoing section, have been applied to this subgroup XII and I—XI taken together. In table 6 we find the results of the F-test:

TABLE 6
F-test for subgroup XII in comparison with the pooled subgroups I–XI:

Sub-groups	Growth after	s^2	ν [6])	F	Bilateral tail-probability
XII I–XI	7 weeks 7 weeks	402 390	17 82	1.04	> 0.40
XII I–XI	12 weeks 12 weeks	801 622	17 44	1.29	0.40

As in none of these two cases the variances differ significantly, STUDENT's test can be applied without objections; table 7 gives the results:

TABLE 7
Student's test for subgroup XII in comparison with the pooled subgroups I–XI:

Sub-groups	Growth after	$\bar{x}$ resp. $\bar{y}$	t	ν [7])	Bilateral tail-probability
XII I–XI	7 weeks 7 weeks	173 157	3.26	99	0.001
XII I–XI	12 weeks 12 weeks	231 218	1.84	61	0.07

[5]), [6]) and [7]) See footnote 4.

We conclude from table 7, that for 7 weeks the means differ significantly whilst for 12 weeks there is a weak indication for a difference between the means.

It does not seem advisable therefore to include subgroup XII in the summerbutter dietgroup. On the other hand we may remark, that this subgroup had a very large increase in weight; thus the exclusion of this group does not lead to flattered results. The conclusions of section 2. 4 are thus certainly not refuted by the data of subgroup XII.

3. Experiments of Groot and Nieman

3. 1. *The diets*

In these experiments the following diets were used:

A: Butter
B_1: Butterfat
B_2: Butterfat + AD [8])
C_1: Butter fatty acids + AD
C_2: Butter fatty acids + OV [9])
C_3: Butter fatty acids + AD + OV
D_1: Arachis oil + OV
D_2: Arachis oil + AD
D_3: Arachis oil + AD + OV
E_1: Arachis oil fatty acids + AD
E_2: Arachis oil fatty acids + OV
E_3: Arachis oil fatty acids + AD + OV.

In these experiments the litters of the rats were known, which made it possible to use more refined techniques of statistical analysis than in the previous case.

3. 2. *Investigation of the association of initial weight and growth*

To test the hypothesis, that the initial weight and the growth of the rats during a given period are stochastically independent, the cornertest of association (see Olmstead and Tukey, 1947) has been applied to the growth on diet D_1 and that on diet D_2, in both cases for 7 weeks. These groups were the largest groups available.

The results of this test are given in table 8:

TABLE 8

Testing the association of the growth and the initial weight by applying the cornertest:

Diet	Growth after	s	Bilateral tail-probalibity
D_1	7 weeks	−3	0.6
D_2	7 ,,	0	1

So there is no reason to assume an association of initial weight and growth.

[8]) AD: vitamins A + D.
[9]) OV: unsaponifiable fraction of butter.

3. 3. *Method of the statistical analysis*

The following method has been used to compare the diets: the difference in growth was calculated for pairs of littermates, the members of which were fed on different diets. Rats which had no littermate in a certain comparison of diets were omitted in that part of the analysis. When two diets have the same nutritive value, the probability of a given positive difference is equal to the probability of an equally large negative difference; i.e. the differences in growth of pairs of littermates are in that case distributed symmetrically with respect to zero. This symmetry has been tested by means of a test for symmetry (test R_2 of J. HEMELRIJK (1950), section 5. 3).

The advantage of this test over that of STUDENT is, that less assumptions need be made. E.g. the assumption of normality is superfluous and the variances of the subgroup need not be equal. From the observations [10]) it is easy to see that in this experiment a number of subgroups differ considerably from the rest and that pooling the subgroups would therefore be inadvisable in this case. The validity of the test for symmetry used is not affected by these differences.

We have compared the following pairs of diets:

a for seven weeks:

1. B_2 and C together with D_2 and E
2. A with D_1
3. A with D_2
4. B_1 with D_2
5. B_2 with D_2
6. C_1 with E_1
7. C_2 with E_2
8. C_3 with E_3
9. D_1 with D_2,

b for twelve weeks:

1. A with D_1
2. A with D_2
3. B_1 with D_2
4. B_2 with D_2
5. D_1 with D_2.

Other comparisons were impossible because either no pairs of littermates could be found or the number of these pairs was too small.

The numbers of pairs of littermates which the subgroups contribute to a given comparison are to be found in table 9 (for 7 weeks) and table 10 (for 12 weeks).

In table 11 a survey of the results of the test for symmetry is given.

---

[10]) See part I, tables 7–22.

TABLE 9

Numbers of pairs of littermates which the subgroups contribute to the comparisons for seven weeks [11])

Comparison	Diets compared	Subgroups																Total
		AR	*AW*	*AX*	*BB*	*BJ*	*BS*	*BU*	*BW*	*BX*	*CA*	*CC*	*CK*	*DC*	*DG*	*DM*	*DP*	
a_1 [12])	$B_2 + C - (D_2 + E)$	–	–	–	–	–	–	–	8	–	10	–	–	7	18	6	14	63
a_2	$A - D_1$	5	2	8	6	4	–	–	–	–	–	–	–	–	–	–	–	25
a_3	$A - D_2$	5	3	5	–	–	–	–	–	–	–	–	–	–	–	–	–	13
a_4	$B_1 - D_2$	–	–	–	–	–	4	7	–	7	–	3	9	–	–	–	–	30
a_5	$B_2 - D_2$	–	–	–	–	–	–	–	8	–	10	–	–	–	–	–	–	18
a_6	$C_1 - E_1$	–	–	–	–	–	–	–	–	–	–	–	–	7	9	6	7	29
a_7	$C_2 - E_2$	–	–	–	–	–	–	–	–	–	–	–	–	–	9	–	–	9
a_8	$C_3 - E_3$	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	7	7
a_9	$D_1 - D_2$	7	1	7	–	–	6	–	–	–	–	–	–	–	–	–	–	21

TABLE 10

Numbers of pairs of littermates which the subgroups contribute to the comparisons for twelve weeks [13])

Comparison	Diets compared	Subgroups								Total
		AR	*AW*	*BB*	*BJ*	*BS*	*BU*	*BW*	*CC*	
b_1	A–D_1	4	2	5	4	—	—	—	—	15
b_2	A–D_2	4	3	—	—	—	—	—	—	7
b_3	B_1–D_2	—	—	—	—	4	7	—	3	14
b_4	B_2–D_2	—	—	—	—	—	—	8	—	8
b_5	D_1–D_2	5	1	—	—	7	—	—	—	13

TABLE 11

Symmetry test R_2 for a number of pairs of groups:

Comparison	Diets compared	n [14])	r	u	v	Bilateral tail-probability
a_1	$B_2 + C - (D_2 + E)$	63	33	29	24	$< 10^{-3}$
a_2	$A - D_1$	25	13	10	4	0.053
a_3	$A - D_2$	13	7	7	4	0.007
a_4	$B_1 - D_2$	30	16	14	7	0.006
a_5	$B_2 - D_2$	18	11	11	5	$< 10^{-3}$
a_6	$C_1 - E_1$	29	16	13	8	0.018
a_7	$C_2 - E_2$	9	5	5	4	0.004
a_8	$C_3 - E_3$	7	4	4	3	0.015
a_9	$D_1 - D_2$	21	11	9	6	0.053
b_1	$A - D_1$	15	8	7	4	0.07
b_2	$A - D_2$	7	4	4	3	0.015
b_3	$B_1 - D_2$	14	7	7	4	0.009
b_4	$B_2 - D_2$	8	4	4	3	0.04
b_5	$D_1 - D_2$	13	7	6	3	0.09

[11]) See part I, tables 7–22.

[12]) a_1 is the same as a_5, a_6, a_7 and a_8 pooled.

[13]) See part I, tables 7, 8, 10–14 and 17.

[14]) n is the number of differences, $u + v$ is the number of positive differences; the meaning of r, u and v is explained in HEMELRIJK (1950).

From table 11 we conclude that for all comparisons, except a_2, a_9, b_1 and b_5 (where the result is not unambiguous), there is a difference in nutritive value between the compared diets. In all cases the first mentioned diet has given the larger increase in weight.

3. 4. *A complication*

Although the comparisons a_2, a_9, b_1 and b_5 of table 11 give a small tail-probability it is not small enough to warrant the conclusion, that a difference between the corresponding diets exists.

During the statistical analysis we found that (at least for 7 weeks) this was possibly caused by the fact that in one subgroup with diet D_1 (i.e. in subgroup *AX* [15])) the growth of the rats was much larger than in the other subgroups with the same diet. The question now arises whether this was caused by exceptional circumstances.

This was tested in two ways:

I. *First method:* There are 5 subgroups [16]), of which *AX* is one, which may be used to compare the diets *A* and D_1. Every one of these subgroups gave a number of growth-differences of pairs of littermates. By means of a 2×2-table, the hypothesis was tested, that the probability of a negative difference was the same for subgroup *AX* as for the other subgroups taken together. The 2×2-table is given in table 12. We used the exact method of R. A. FISHER (see R. A. FISHER, 1950, p. 96), computing the bilateral tail-probability as the sum of the probabilities of all possible results, which are not more probable than the actual result of the experiment.

TABLE 12

Number of positive and negative differences in growth, by comparing the diets A and D_1, in the subgroup AX and the other subgroups with the same diets:

Subgroups	Number of positive differences	Number of negative differences	Total
AR	3	2	5
AW	2	0	2
BB	3	3	6
BJ	4	0	4
sub-total	12	5	17
AX	1	7	8
total	13	12	25

The bilateral tail-probability proved to be 0.01. This has to be multiplied by 5, to take the fact into account, that subgroup *AX* was chosen from the 5 subgroups as the group with the smallest number of positive differences.

[15]) See part I, table 9.

[16]) See part I, tables 7, 8, 9, 10 and 11.

II. *Second method:* We also tested, whether the growth on diet D_1 in subgroup AX is larger than the growth on the same diet in the other subgroups, during the same season (these were the subgroups AR and AW). For this purpose WILCOXON's two-sample test (see WILCOXON 1945 and MANN and WHITNEY 1947) was applied to the observations of subgroup AX in comparison with all observation of the pooled subgroup AR and AW.

TABLE 13

Comparison of subgroup AX with diet D_1 with the other subgroups with the same diets by applying Wilcoxon's two sample test:

n [17])	m [18])	U	$\frac{d}{\sigma}$	Bilateral tail-probability
24	15	288	3.1	0.002

In this case the tail-probability has to be multiplied by 3, to take into account the fact that subgroup AX was chosen from three subgroups as the group with the larger average growth.

From the tables 12 and 13 we may conclude that the growth in subgroup AX was indeed abnormally high in consequence of unknown special ircumstances. This makes it preferable to leave this subgroup out of consideration.

3. 5. *Final comparison of the diets A, D_1 and D_2*

On account of the result of section 3. 4 subgroup AX with diet D_1 is now left out of consideration and for the comparisons a_2 and b_9 the symmetrytest R_2 is applied once more. The results are given in table 14:

TABLE 14

Symmetry test R_2 applied to two pairs of groups:

Comparison	Diets	n	r	u	v	Bilateral tail-probability
a_2	$A-D_1$	17	9	9	2	0.004
a_9	D_1-D_2	14	8	6	2	$>$ 0.10

As parameterfree methods like this test for symmetry are sometimes less powerfull than methods based on special assumptions as e.g., the assumption of normality and of equal variances, we have for those comparisons which gave no significance (i.e. for a_2, a_9, b_1 and b_5) applied STUDENT's test for the mean of a normal distribution (see e.g. M. G. KENDALL 1947, vol. II p. 98) on the same differences where before we applied the symmetry test.

With STUDENT's test we test the hypothesis that the mean of the distribution of the differences in growth is zero and the test is based on the

17) Number of observation in subgroup AX (see table 9, part I).

18) Number of observations in the pooled subgroups AR and AW (see tables 7 and 8, part I).

ratio of the mean and the standard deviation of the sample. In table 15 we find the results, while table 16 gives a summary of the bilateral tail-probabilities which are found with the symmetry test and Student's test for the comparisons a_2, a_9, b_1 and b_5.

TABLE 15

Student's test:

Comparison	Diets	$\bar{x}$	t	ν [19])	Bilateral tail-probability
a_2	$A-D_1$	19	3.3	16	0.004
a_9	D_1-D_2	7	1.0	13	0.34
b_1	$A-D_1$	20	2.8	14	0.014
b_5	D_1-D_2	22	2.2	12	0.046

TABLE 16

Summary of the tail-probabilities found with the symmetry test and Student's test for the comparisons a_2, a_9, b_1 and b_5 with and without subgroup AX:

Comparison	Diets	With subgroup AX	Without subgroup AX	
		Symmetry test		Student
a_2	$A-D_1$	0.053	0.004	0.004
a_9	D_1-D_2	0.053	$>$ 0.10	0.34
b_1	$A-D_1$	—	0.07	0.014
b_5	D_1-D_2	—	0.09	0.046

We see that the tail-propability obtained by applying STUDENT's test are indeed sometimes smaller than those found with the symmetry test.

As from § 3. 4 it is justified not to consider subgroup AX while the use of STUDENT's test alone cannot be considered as justified, we conclude from the experiments of GROOT and NIEMAN that a difference in nutritive value exists between:

1. Butterfat, butter fatty acids on the one hand and arachis oil, arachis oil fatty acids on the other hand, where to both diets AD and/or OV is added (see a_1; table 11).
2. Butter and arachis oil + OV (see a_2 and b_1; table 16).
3. Butter and arachis oil + AD (see a_3 and b_2; table 11).
4. Butterfat and arachis oil + AD (see a_4 and b_3; table 11).
5. Butterfat + AD and arachis oil + AD (see a_5 and b_4; table 11).
6. Butter fatty acids + AD and arachis oil fatty acids + AD (see a_6; table 11).
7. Butter fatty acids + OV and arachis oil fatty acids + OV (see a_7; table 11).
8. Butter fatty acids + AD + OV and arachis oil fatty acids + AD + OV (see a_8; table 11).

In all these cases the first mentioned diet caused the larger increase in weight.

[19]) See footnote 4.

There is no decisive indication for a difference in nutritive value between arachis oil + OV and arachis oil + AD (see a_9 and b_5; table 16).

I want to thank Prof. Dr J. HEMELRIJK for his advice and criticism which have been very helpful.

LITERATURE

BARTLETT, M. S., The problem in statistics of testing several variances, Proc. Camb. Phil. Soc. **30**, 164 (1934).

DIXON, W. J. and F. J. MASSEY, An introduction to statistical analysis (New York, 1951).

FISHER, R. A., Statistical methods for research workers (London, 1950).

GEARY, R. C. and E. S. PEARSON, Tests of normality, Biometrika office (London, 1938).

HARTLEY, H. O., Testing the homogeneity of a set of estimated variances, Biometrika 31, 249 (1940).

——— and E. S. PEARSON, Tables for testing the homogeneity of a set of estimated variances, Biometrika 33, 296 (1946).

HEMELRIJK, J., A family of parameterfree tests for symmetry with respect to a given point, I and II, Proc. Kon. Ned. Akad. v. Wet. **53**, 945–955 and 1186–1198 (1950).

HOEL, P. G., Introduction to mathematical statistics (London, 1948).

KENDALL, M. G., The advanced theory of statistics (London, 1947).

MANN, H. B. and W. R. WHITNEY, On a test whether one of two random variables is stochastically larger than the other, Annals of mathematical statistics 18, 50 (1947).

OLMSTEAD, P. S. and J. W. TUKEY, A corner test for association, Annals of mathematical statistics **18**, 495 (1947).

WILCOXON, F., Individual comparisons by ranking methods, Biometrics **1** (1945).